Explainable Artificial Intelligence (XAI)

Year: 2023 | Citations: 746 | Authors: Ranu Sewada, Ashwani Jangid, Piyush Kumar, Neha Mishra

---

Abstract

Explainable Artificial Intelligence (XAI) has emerged as a critical facet in the realm of machine learning and artificial intelligence, responding to the increasing complexity of models, particularly deep neural networks, and the subsequent need for transparent decision making processes. This research paper delves into the essence of XAI, unraveling its significance across diverse domains such as healthcare, finance, and criminal justice. As a countermeasure to the opacity of intricate models, the paper explores various XAI methods and techniques, including LIME and SHAP, weighing their interpretability against computational efficiency and accuracy. Through an examination of real-world applications, the research elucidates how XAI not only enhances decision-making processes but also influences user trust and acceptance in AI systems. However, the paper also scrutinizes the delicate balance between interpretability and performance, shedding light on instances where the pursuit of accuracy may compromise explain-ability. Additionally, it navigates through the current challenges and limitations in XAI, the regulatory landscape surrounding AI explain-ability, and offers insights into future trends and directions, fostering a comprehensive understanding of XAI's present state and future potential.